

# Bus stop data quality in OpenStreetMap

JENS FLEMMING\*

April 5, 2024

**Key words:** OpenStreetMap, bus stop, tagging schemes

OpenStreetMap (OSM) is a large open database for geographic data created and maintained by volunteers, see Wikipedia on OSM for an overview and some background. OSM's main data use is rendering an extremely detailed map of the world.

Data quality is an important issue for applications like routing of pedestrians to public transport facilities. In this report we describe different schemes for mapping bus stops in OSM and we provide statistics on usage of those schemes, the good ones and the not so good ones.

## OSM bus stop basics

OSM does not have one fixed scheme for mapping bus stops. Instead, there exist several tags for describing bus stops or parts of bus stops. Tag usage changes over time. The decision for a tagging scheme usually is based on

- what other mappers have done so far,
- what the OSM wiki currently suggests,
- what applications (renderers, routers,...) do process.

Tagging schemes should meet all three points. The first point (other mappers) is the main concern of this report. Addressing this point requires a deep look into the OSM database. Before we come to the statistics part, we summarize what OSM wiki has about bus stops. The question which tags applications do process cannot be answered in general. In this report we focus on OSM's standard style Carto for the Mapnik renderer.

A bus stop in the OSM data base consists of one of the following collections of objects:

- one node on the road or beside the road,

---

\*Zwickau University of Applied Sciences, Faculty of Physical Engineering/Computer Sciences, D-08012 Zwickau, Germany, jens.flemming@fh-zwickau.de.

- two nodes, one on the road, one beside the road,
- a line or an area beside the road,
- one node on the road, a line or an area beside the road,
- anything else someone finds useful.

Next to many individual mapping schemes there exist several documented ones:

- the original mapping scheme, sometimes referred to as PTV1 (PT = public transport),
- the approved PTV2 proposal from 2011,
- the not yet approved PTV3 proposal from 2018 (discussion is still active),
- the not so well documented behavior of Carto.

The original mapping scheme has turned out to be not expressive enough. PTV2 introduced a rather powerful (and complex!) tagging scheme. PTV3 tries to combine the power of PTV2 with the simplicity of PTV1. The Carto style mainly follows PTV1.

We briefly summarize the bus related relevant parts of PTV1, PTV2 and Carto rendering below. Since it's not yet approved, we do not discuss PTV3 here.

## PTV1

Following PTV1 bus stops are nodes on the road or beside the road with the `highway=bus_stop` tag. The node should be placed beside the road to indicate the passenger waiting area. For simplicity and for making OSM data more accessible to routing software some mappers place the node on the road.

There's also a `highway=platform` tag to map the waiting area as a line, a rectangle or some other shape.

## PTV2

In PTV2 a bus stop usually is a group of two objects: a stopping position on the road and a platform beside the road. PTV2 also knows stop areas (relations combining all objects of a stop) and routes (lists of stops for describing bus routes).

The stopping position is a node with `public_transport=stop_position` and `bus=yes`. A node on the road with `highway=bus_stop` is interpreted as stop position, too.

A platform is a node or a way or an area (closed way with `area=yes`) with `public_transport=platform`. Often there's also a `bus=yes` to indicate the purpose of the platform, but that's not required by PTV2. A node beside the road with `highway=bus_stop` is interpreted as platform, too.

Routes in PTV2 should contain both stopping positions and platforms.

## Carto rendering

The standard OSM map only renders the following bus stop related objects:

- nodes with `highway=bus_stop` (blue bus stop sign),
- areas with `highway=bus_stop` (blue bus stop sign at the area's center),
- ways/areas with `highway=platform` (gray line/area).

Carto does not use the PTV2 tags for rendering because there are too many ambiguous tag combinations in the OSM database. It's a long and ongoing discussion:

- Add rendering for `public_transport=platform`
- Render public transport platforms and stops (new scheme, relation-free implementation for now)
- many threads in OSM community forum

## Mapping schemes

Based on PTV2 and Carto's behavior we want to derive bus stop tagging schemes which on the one hand are compatible with PTV2 and on the other hand produce acceptable renderings with Carto. We distinguish between stops without platform (just a pole on the side walk, for instance) and stops with platform (waiting area is not part of a sidewalk and has been solely constructed for boarding the bus).

Each tagging scheme gets an abbreviation, which will be used in the bus stop statistics below.

### Stops without platform

**Stopping position only (SPN)** The simplest bus stop in OSM is

- a node on the road with `highway=bus_stop`.

That's the PTV1 way for tagging a bus stop (although placing the node beside the street is preferred). PTV2 interprets the node as a stopping position. Stopping positions without platform node/way/area are okay in PTV2 (mapping the platform, too, is preferred). Carto will show a bus stop sign on the road.

**Platform node only (PFN)** Another simple bus stop is

- a node beside the road with `highway=bus_stop`.

That's the preferred PTV1 way for tagging a bus stop. PTV2 interprets the node as a platform. Platforms without stopping position seem to be okay in PTV2 (mapping the stopping position, too, is preferred). Carto will show a bus stop sign beside the road.

**Two nodes (PFN+SPN)** The preferred PTV2 way for tagging a bus stop without platform is

- a node on the road with `public_transport=stop_position`,
- a node beside the road with `public_transport=platform`.

Both nodes should also have `bus=yes`.

Carto won't show any of the two nodes. To get correct rendering with Carto one of the nodes should additionally have `highway=bus_stop`, preferably the platform node. Then Carto shows the bus stop sign on the map.

## Stops with platform

**Platform and stopping position (PFW+SPN)** The PTV2 way for mapping a stop with platform is

- a node on the road with `public_transport=stop_position`,
- a way/area with `public_transport=platform`.

Both objects should also have `bus=yes`.

None of both objects gets rendered by Carto. To get a bus stop sign on the road add `highway=bus_stop` to the stopping position node. To get the platform rendered add `highway=platform` to the platform way/area.

In case of a platform area one may make Carto render a bus stop sign at the platform by adding `highway=bus_stop` to the area. But then the platform itself won't get rendered anymore.

**Platform with two nodes (PFW+PFN+SPN)** To get better rendering with Carto one may use three objects for a bus stop:

- a node on the road with `public_transport=stop_position`,
- a way/area with `public_transport=platform`,
- a node on the platform with `public_transport=platform` (may be a node of the platform way/area).

All three objects should also have `bus=yes`.

None of the three objects gets rendered by Carto. To get a bus stop sign on the platform add `highway=bus_stop` to the platform node. To get the platform rendered add `highway=platform` to the platform way/area.

This tagging scheme might contradict the OSM rules 'one OSM object per real object' and 'no mapping for the renderer'. But we could argue that the platform node refers to the bus stop pole and the platform way/area to the platform.

Another issue with this tagging scheme is that PTV2 on the one hand expects that a route contains all available stop positions and platforms. On the other hand, PTV2 asserts that each stop is included in routes with two elements, the stop position and the platform. The situation that the platform is represented

by two objects is not covered by PTV2 (this fact has been pointed out to the author by Toni Erdmann in a discussion in OSM community forum). If this tagging scheme is used, only one of the two platform objects should be added to routes (else, routing software might run into troubles). Additional tags like `bench`, `bin`, `wheelchair` aso. should then go to the platform object included into routes.

### **Platform with two nodes, PTV2 compatible (PFW+PFN+SPN+v2)**

With a slight modification the previous tagging scheme becomes fully PTV2 compatible:

- a node on the road with `public_transport=stop_position`,
- a way/area with `highway=platform`,
- a node on the platform with `public_transport=platform` (may be a node of the platform way/area).

Here the platform way/area has no `public_transport=*` tag. Thus, there's only one platform object (the node), which in addition gets `highway=bus_stop` to make Carto show the bus stop sign. The platform node represents the stop including all relevant information, whereas the platform way/node plays the role of the ground the stop rests on. This is in complete analogy to the two-nodes variant PFN+SPN for stops without platform, where the stop rests on a sidewalk and the sidewalk does not have any public transport related tags.

## **Scheme usage statistics**

The decision for a tagging scheme to use for adding bus stops to the OSM database should also be based on what schemes other mappers have been following so far in the region of interest. Due to the multitude of possible object and tag combinations it's hard to get an overview of scheme usage. In this section we provide and use some code to automatically identify bus stop tagging schemes for a region of interest.

## **Code download and usage**

Here is the Python code in a Jupyter notebook:

- Download `bus_stops.ipynb` (if your browser opens the text file instead of downloading, click 'Save as' or copy the text into a text editor and then save it).

The code is of poor quality. Don't expect a piece of highly polished software! It's more like a proof of concept.

Usage:

0. Install Jupyter Lab and Python packages `folium`, `geopy`, `numpy`, `osmium`.

1. Get OSM data of your region of interest (`*.osm` or `*.osm.pbf` format). Use JOSM or OSM's export feature or BBBike file list or BBBike extract feature.
2. Write your data's file name to the first cell in the notebook.
3. Run all cells in the notebook (may take a while).
4. The last cell in the notebook shows the interactive map. In addition, a `*.html` file with the map appears in your working directory.

In the upper right corner of the map you may choose what is shown on the map. Hovering nodes you get some information about the node. Following abbreviations are used:

- `has_ptpf` = has `public_transport=platform` and `bus=yes`,
- `has_ptsp` = has `public_transport=stop_position` and `bus=yes`,
- `has_hwbs` = has `highway=bus_stop`,
- `has_hwpf` = has `highway=platform`,
- `in_route` = is member of some route,
- `on_road` = is on a road (not beside road),
- `closest_sp_node` = ID of closest stop position node (max. 30 meters distance),
- `closest_pf_node` = ID of closest platform node (max. 30 meters distance).

## Results for selected regions

We use the scheme abbreviations introduced above. An `s` appended to a component of the abbreviation means that a bus stop sign is shown at corresponding location. Without `s` no bus stop sign is shown. We only consider combinations the author considers acceptable tagging practice:

scheme	acceptable	comment
SPN	yes	bus stop on road, no rendering
SPNs	yes	bus stop on road with rendering
PFN	yes	bus stop beside road, no rendering
PFNs	yes	bus stop beside road with rendering
PFN+SPN	yes	preferred PTV2 bus stop, no rendering
PFNs+SPN	yes	preferred PTV2 bus stop with bus stop sign at platform
PFN+SPNs	no	as before, but bus stop sign on road although there's a platform
PFNs+SPNs	no	as before, but with two bus stop signs
PFW+SPN	yes	PTV2 platform way/area, platform may be rendered, but no bus stop sign
PFWs+SPN	yes	PTV2 platform area with bus stop sign at platform
PFW+SPNs	yes	PTV2 platform way/area with bus stop sign on road

scheme	acceptable	comment
PFWs+SPNs	no	PTv2 platform area with two bus stop signs
PFW+PFNs+SPN	yes	platform way/area with bus stop sign at platform
PFWs+PFNs+SPN	no	platform area with two bus stop signs at platform
PFW+PFNs+SPNs	no	platform way/area with two bus stop signs
PFWs+PFNs+SPNs	no	platform area with three bus stop signs
PFW*+PFN+SPN*	no	useless, because the platform node's purpose is to have a bus stop sign
PFW+PFNs+SPN+v2	yes	platform way/area with bus stop sign at platform
PFWs+PFNs+SPN+v2	no	not possible because <code>highway=*</code> would have two values
PFW+PFNs+SPNs+v2	no	platform way/area with two bus stop signs
PFWs+PFNs+SPNs+v2	no	not possible because <code>highway=*</code> would have two values
PFW*+PFN+SPN*+v2	no	useless, because the platform node's purpose is to have a bus stop sign

Route membership is relevant for categorization, too. If something is wrong with route membership (only one of the two nodes in PFN+SPN belongs to a route, for instance), then the bus stop does not belong to the category.

Regions shown here have been chosen without thinking to much about it. One is the author's home region. The others are randomly chosen cities and regions in Germany.

Stops without platform way/area:

scheme	Chemnitz-				
	Berlin	Zwickau	Darmstadt	Munich	Reutlingen
SPN	180 ( <b>61.9%</b> )	176 (4.7%)	25 (2.0%)	33 (2.2%)	58 (3.3%)
SPNs	1 (0.3%)	595 ( <b>15.8%</b> )	33 (2.6%)	88 (5.8%)	9 (0.5%)
PFN	0	6 (0.2%)	0	0	2 (0.1%)
PFNs	11 (3.8%)	251 (6.7%)	441 ( <b>34.4%</b> )	117 (7.7%)	76 (4.3%)
PFN+SPN	0	14 (0.4%)	0	0	5 (0.3%)
PFNs+SPN	99 ( <b>34.0%</b> )	2353 ( <b>62.4%</b> )	773 ( <b>60.3%</b> )	1128 ( <b>74.1%</b> )	1579 ( <b>89.7%</b> )

scheme	Chemnitz-				
	Berlin	Zwickau	Darmstadt	Munich	Reutlingen
other	0	375 (9.9%)	9 (0.7%)	156 (10.2%)	32 (1.8%)
total	291	3770	1281	1522	1761

Stops with platform way/area:

scheme	Chemnitz-				
	Berlin	Zwickau	Darmstadt	Munich	Reutlingen
PFW+SPN	12 (0.4%)	21 (14.5%)	57 ( <b>18.1%</b> )	62 (2.7%)	19 (11.2%)
PFWs+SPN	0	0	5 (1.6%)	361 ( <b>15.8%</b> )	8 (4.7%)
PFW+SPNs	1 (0.0%)	50 ( <b>34.5%</b> )	127 ( <b>40.3%</b> )	1631 ( <b>71.3%</b> )	4 (2.4%)
PFW+PFNs +SPN	3288 ( <b>97.6%</b> )	28 ( <b>19.3%</b> )	54 ( <b>17.1%</b> )	132 (5.8%)	97 ( <b>57.4%</b> )
PFW+PFNs ' +SPN+v2	0	1 (0.7%)	0	2 (0.1%)	1 (0.6%)
other	67 (2.0%)	45 (31.0%)	72 (22.9%)	100 (4.4%)	40 (23.7%)
total	3368	145	315	2288	169

Maps with categorized stops (large ones may significantly slow down your browser):

- Berlin (18 MB)
- Chemnitz-Zwickau (6 MB)
- Darmstadt (5 MB)
- Munich (25 MB)
- Reutlingen (5 MB)

Results for the five regions are quite different. Berlin uses PFW+PFNs+SPN for almost all stops. Munich is quite uniform, too, with PFNs+SPN and PFW+SPNs for three quarters of all stops. Reutlingen's stops without platform way/area mostly follow PFNs+SPN like in Munich. Scheme usage in Chemnitz-Zwickau region and in Darmstadt is far from being consistent. Especially, one third of all stops with platform way/area in those regions (and Reutlingen) do not follow one of the acceptable schemes.

From manual spot-checking one sees that a major issue with 'other' platform way/area stops is that only the stop position is in the routes (not allowed in PTV2). For stops without platform way/area the most common deviation from the schemes listed above are additional bus stop signs on the road although there's a stop bus sign at the platform already.



Categorization of stops heavily relies on the automatic assignment of stop positions to platform nodes/ways/areas and of platform nodes to platform ways/area. The assignment procedure uses simple distance measuring (closest node within 30 meter radius around node or way/area). There seem to be few cases where this algorithm yields incorrect results, mostly because the stop position is too far away from the platform or the closest stop position belongs to another platform.

Another issue with the approach are bus platforms having `public_transport=platform` but no `bus=yes`. In view of PTV2 this is okay. But our code does not recognize such platforms as bus related. Consequently, there will be stop positions without platform. Not expecting `bus=yes` for platforms would bring many tram-only platforms into the game, again causing wrong categorization.

Starting bus stop mapping, the author came up with the idea to use `PFW+PFNs+SPN+v2`, which seemed to be an elegant tagging scheme meeting the requirements of PTV1, PTV2 and Carto and which by the way establishes stronger parallels between stops with and without platform ways/areas (cf. description of scheme above). Based on a discussion in OSM community forum `PFW+PFNs+SPN+v2` seemed to be quite uncommon in mapping practice. That discussion has been the starting point for the bus stop statistics project. To the author's regret, `PFW+PFNs+SPN+v2` indeed is rarely used in OSM.